DOCUMENT RESUME

ED 442 820 TM 031 240

AUTHOR Nasser, Fadia
TITLE Understanding Students' Statistical Thinking: An Error Analysis Approach.
PUB DATE 2000-04-00
NOTE 32p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Concept Formation; Females; Foreign Countries; Higher Education; Learning; Multiple Choice Tests; Research Methodology; *Statistics; Student Attitudes; *Undergraduate Students
IDENTIFIERS *Errors; Israel

ABSTRACT

This exploratory study examined errors that students commit in solving multiple-choice questions about descriptive statistics and basic concepts in research methods. The sample consisted of 81 undergraduate students in an introductory statistics course. The majority (85%) of the participants were female students enrolled in education, sociology, psychology, and arts programs at a major university in Israel. The results indicate that the most frequently detected errors were confusing concepts, misinterpreting descriptive information, applying inappropriate procedures, and applying partial information. Analyses reveal that potential sources of students' errors include assimilation of statistical concepts into inappropriate schema, failure to use knowledge sources, and lack of ability to relate and combine knowledge from different sources. Suggestions for further research are also provided. (Contains 1 figure, 1 table, and 22 references.) (Author/SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 442 820



# Understanding Students' Statistical Thinking: An Error Analysis Approach

Fadia Nasser

Tel Aviv University

F. Nasser

2B

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper Presented at the Annual Meeting of the American Educational Research Association

New Orleans, LA, April 24-28, 2000

TM031240



BEST COPY AVAILABLE

## Abstract

This exploratory study examined errors that students committed in solving multiple-choice questions about descriptive statistics and basic concepts in research methods. The sample consisted of 81 undergraduate students in an introductory statistics course. The majority (85%) of the participants were female students enrolled in education, sociology, psychology, and arts programs at a major university in Israel. The results indicated that the most frequently detected errors were confusing concepts, misinterpreting descriptive information, applying inappropriate procedures, and applying partial information. Analysis reveal potential sources of students' errors include assimilation of statistical concepts into inappropriate schemata, failure to use knowledge sources, and lack of ability to relate and combining knowledge from different sources. Suggestions for further research are also provided.

Introductory statistics is a prerequisite for attaining a degree in most disciplines. The inclusion of statistics in most academic programs reflects the importance that policy makers attribute to this subject. Yet, the difficulty that introductory statistics courses impose on students and the low levels of achievement in statistics are widely reported (Auizmendi, 1991; Del Vecchio, 1995; Glencross & Cherian 1992; Thompson & Smith 1982; Vidal-Madjar, 1978; Watts, 1991).

Researchers have used different approaches in attempting to understand students' achievement in statistics. Students' performance has been studied by examining the effects of cognitive factors, such as mathematics ability and background, and affective factors, such as mathematics anxiety, attitudes towards statistics, statistics anxiety, and students' motivation on statistical achievement (Feinberg & Halprin, 1978; Gal & Ginsburg, 1994; Nasser, 1998, 1999; Wisenbaker & Scott, 1997). The information concerning cognitive and affective factors and the way they affect students' performance in statistics is critical because the information regarding these factors is essential for planning and executing effective instruction.

Another approach to understanding students' performance is based on analyzing the errors these students commit. This approach has been widely used in mathematics (Babbitt, 1990; Jitendra & Kameenui, 1993; Nasser, 1989) and has recently been adopted by some researchers in the area of probability and statistics (Cohen, Smith, Chehile, Burn, & Tsai, 1996; Herman, 1997; O'Connell, 1997). In the context of mathematics, researchers have extensively studied students' error patterns in solving mathematical problems in order to understand students' reasoning processes and to provide instruction for remedying errors and misconceptions (Brown & Burton, 1978; Nasser, 1989; Vikali, 1985). One aspect of error analysis that has been reported in the literature involves classifying errors into categories. These categories range from simple

computational errors (arithmetic errors) to more complex errors derived from theories grounded in misuse of reading and information processing (Jitendra & Kameenui, 1993). For example, Babbit (1990) classified errors that students made in dealing with mathematical tasks into computational errors, operational errors, non-attempt errors, and errors that cannot be classified with the previous categories.

Few studies have dealt with error analysis in the area of probability and statistics (Mevarech, 1983; O'Connell, 1993). For example, O'Connell (1993) classified students' errors in probability problem solving into four categories, which he referred to as text comprehension, conceptual, procedural, and arithmetic. Herman (1997) investigated statistical content errors made by students in an educational psychology course when answering multiple-choice questions. Consequently, he suggested that item analysis findings from multiple-choice examinations could be used to reveal students' conceptual misunderstandings, improve classroom instruction, and refine test-item writing.

In order to provide effective remedies for students' errors, it is imperative to identify their sources and to determine how these sources translate into errors. In other words, it is very essential to understand students' reasoning and the cognitive processes involved in solving statistical tasks. To this end, researchers have attributed students' errors and misconceptions to the use of irrelevant schemes, to the use of partial information, and to the conflict between the intuitive and scientific meanings of statistical concepts (Mevarech, 1983, Shaughnessey, 1993).

Two noteworthy points emerged from reviewing the literature pertaining to error analysis. First, very few studies have been conducted in this area. Furthermore, most of these studies were conducted at lower grade levels, particularly with elementary school children. Second, the limited studies conducted with students in higher education, especially with those who are not

mathematically oriented, are insufficient for studying important aspects of students' errors. Therefore, further research is called for in order to expand the body of knowledge concerning students' errors and the thinking processes responsible for their occurrence.

The purpose of this study was to attempt to provide some remedies for this situation. Specifically, the study aimed to: (a) explore errors that students make in solving multiple-choice questions about descriptive statistics and basic concepts in research methods, and (b) identify and describe possible sources of students' errors in order to gain a better understanding of the possible cognitive and reasoning processes involved.

## Method

### Sample

The sample consisted of 81 undergraduate students in an introductory statistics course. The majority (85%) of the participants were female students enrolled in education, sociology, psychology, and the arts at a major university in Israel. Introductory statistics was a required course for all the participants.

### Instruments

An achievement test was administered at the end of the semester. The test included 50 questions of which 39 were multiple-choice; the remaining required short answers. The test included questions about basic concepts in statistics and research methods, and descriptive statistics. This study focused only on the multiple-choice questions: basic concepts in statistics and research methods (13 questions), distributions and graphs (5 questions), measures of central tendency (4 questions), measures of dispersion (8 questions), measures of relative location (4 questions), and the normal distribution (5 questions). The internal consistency associated with the test as indicated by Croncach-$\alpha$ is .82.

BEST COPY AVAILABLE

To facilitate the diagnosis of errors and misconceptions frequently made by students, the distracters in the multiple-choice questions were drawn from common errors made by students from the same population. An example of a question on the test was:

*The range of scores on a cognitive ability test is 300 to 600. If these scores are normally distributed, then the value of the standard deviation was:*

*(a) 50 (b) 1/2 (c) 2 (d) 300*

The selection of one of the three distracters (b, c, and d) implies a lack of understanding of the relationship between the range and the standard deviation of the normal distribution (the range is about six times the S.D.) Yet, an erroneous answer may be a result of unsuccessful guessing.

## Results

In order to obtain a comprehensive picture of students' performance on the achievement test, their answers were explored using item analysis. Furthermore, item difficulty as indicated by the percentage of students who correctly answered each question and the frequency of errors made within each statistical topic were obtained. Table 1 provides a summary description of the topics and tasks on the achievement test, the frequent errors made by students, and the item difficulty indices.

Insert Table 1 Here

As shown in Table 1, all items, except one (item 17), had moderate to high item-difficulty indices indicating that the test, in general, was not difficult. Concurrently, several types of errors emerged from the analysis of the erroneous answers. The most frequent errors were: confusing concepts, misinterpreting descriptive information, and applying inappropriate procedures. Few computational errors were found.

The number of errors within each of the studied topics is displayed in Figure 1. The mean number of errors per question was used as a measure of the number of errors per topic. As indicated by Figure 1, the largest mean number of errors per question associated with questions about distributions and graphs (21 errors/question). At the same time, the number of errors per question within the remaining topics (basic concepts in descriptive statistics and research methods, measures of relative location, measures of dispersion, normal distribution and measures of central tendency) ranged from 14-16.

Insert Figure 1 Here

In order to illuminate potential sources of frequent errors and misconceptions, it is imperative to define what knowledge and cognitive skills students should possess to successfully solve statistical questions. Stated differently, the knowledge and cognitive skills necessary for correctly answering each question on the achievement test should be defined during the planning phase of the item construction-process. A sample of questions and the associated frequent errors are described in the following section.

Task and Error Analysis

A sample of eight of the 39 questions on the achievement test has been selected to illustrate the results concerning tasks and associated frequent errors. These questions were selected because they were found to be relatively difficult and because one of the distracters in each of these eight questions attracted a considerable proportion of the students. Such attractive distracters reflect a problem on the part of the student and/or the test designer; hence, they are worthy of special treatment. It should be noted, that the selected questions represent all the topics included in the

test.

**Q1. Basic Concepts in Statistics and Research Methods (Variables and Hypotheses)**

*A group of researchers at the university argue that the scores on the Entrance and the matriculation examinations do not predict the degree of students' success in the first academic year. This argument includes:*

*1. One dependent (criterion) and two independent (predictor) variables*

*2. Two dependent and one independent variables*

*3. One dependent, one independent, and one mediator variable*

*4. This argument is not scientific*

To answer this question correctly, students' should understand the definition (conceptual meaning) of a scientific argument, a hypothesis, and a variable, and be able to distinguish among them. They also should be able to identify the dependent and independent variables in a given situation. Among the participants, 18.5% selected option 4 as an answer. Several possible reasons for this response emerged from discussions with students. For example, some students reported difficulty in understanding the task while others reported difficulty in dealing with arguments that include more than two variables. Students who erred may lack the basic conceptual understanding of scientific arguments, hypotheses, and variables. It is also possible that deficiencies in teaching the topic, such as exposing students to simple arguments and hypotheses that include only two variables, one dependent and one independent, may be responsible for inducing this kind of error. Students who erred on this question were able to recognize that the argument includes more than two variables (they did not select option 2), and that none of these variables is a mediator (they did not select option 3). Alternatively, it is also plausible that the selection of option 4 by some students *is rooted* in knowing that a "scientific argument" is based on empirical data. Because such data were not supplied with the question, they may have considered the argument as unscientific.

However, guessing cannot be excluded as the reason for selecting option 4.

**Q2. Measures of Relative Location**

*Below is information concerning three achievement tests for Debbi's class:*

| | *Bible* | *Mathematics* | *English* |
|---|---|---|---|
| *Mean* | *75* | *75* | *75* |
| *S.D* | *3* | *8* | *5* |

*Debbi's scores in the three subjects are 10 points above the class mean. In which subject is Debbi's achievement relatively the best?*

*1. English 2. Bible 3. Mathematics 4. Her achievement in the three subjects is equal*

To answer this question properly, students' should obtain Debbi's *Z*-score on each one of the tests and interpret the highest *Z*-score to mean the relatively best achievement. Among the students, 44.4% erroneously selected option 3 as an answer. The work papers they handed in along with the test response sheet indicated that some students mistakenly interpreted their mathematically correct answers. Their erroneous interpretation is rooted in confusing the index of variability (C.V.) with the *Z*-score. A potential source of this error is a faulty assimilation of the two measures, that is, the difference between the meanings and applications of each is still vague for a large proportion of the students. It should be clarified that participants in this study had been taught the C.V. index prior to the *Z*-score. Therefore, it seems that although they had no difficulty in successfully obtaining the appropriate measure, they failed to properly interpret their mathematically correct result.

**Q3. Measures of Variability (Interquartiles)**

*A mathematics test that included 150 questions was given to three classes. The results are summarized in Table 1.*

*Table 1*

*Summary of Descriptive Measures for the Math Scores for the Three Classes.*

| | *Class I* | *Class II* | *Class III* |
|---|---|---|---|
| *N* | *24* | *30* | *28* |
| *Mean* | *55* | *90* | *120* |
| *Median* | *45* | *89* | *128* |
| *Q1* | *37* | *82* | *100* |
| *Q3* | *60* | *98* | *138* |
| *S.D* | *Unknown* | *12* | *28* |

*For class II, what is the range of scores obtained by the 15 students at the center of the distribution?*

*1. 82-98 2. 82-89 3. 89-98 4. 78-102*

A correct answer to this question indicates that students: (a) understand the close "the 15 students at the center of the distribution" (50% of N at the center of the distribution), (b) relate this information to the interquartile, (c) obtain Q=Q3-Q1, and (d) select the right answer.

This question turned out to be moderately difficult (difficulty index = 53%). Although the range in option 4 did not exist in Table 1, 20% of the students selected this option as an answer. This in itself supported the possibility that these students did not understand the task and that they probably selected the answer by guessing. Meanwhile, 16% of the participants selected option 3 as an answer. Once again, it seems that these students did not understand the task and/or they lacked a basic conceptual understanding of the meaning of the interquartile. These students associated the

median with the center of the distribution but they were not able to translate the number 15 into 50% of the 30 students in class II. Furthermore, they failed to understand that the scores of these students ranged between Q1 and Q3 rather than between Q3 and the median (Q2). Although, these interpretations of the students' errors are plausible, it is more likely that the students, in general, were unable to comprehend the task or parts of it. This conclusion is supported by the fact that students successfully obtained the correct value of the interquartile when they were required to do so directly in another question on the same test.

**Q4. Distributions, Graphs, and Measures of Central Tendency**

*Based on the information in Table 1, for which class/es is obtaining and reporting the median is more appropriate than the mean?*

*1. Only class I  2. Only class II  3. Only class III  4. Class I and Class III*

To answer this question successfully, students' should be able to (a) detect the shape of the distribution given the information in the table, and (b) select the appropriate measure of central tendency that best characterizes (describes) this distribution.

The difficulty level of this question was 37.4%. Two of the three distracters (options 2 and 3) attracted 22% and 31% of the students, respectively. The selection of these options indicates that students used partial information to obtain their answers. It is unclear whether students used partial information because they were unable to detect the shape of the distribution and/or because they were unable to select a measure of central tendency that appropriately characterizes the distribution. To resolve this confoundedness, it might be helpful to begin with a question in which students are required to determine the shape of the distribution for each of the three classes before proceeding with questions in which they are required to select an appropriate measure of central tendency to characterize the distribution.

**Q5 Normal and Standard Normal Distributions**

*The standard deviation of a normal distribution is:*

*1. Zero 2. 1 3. A positive number 4. Negative, positive numbers or zero*

In order to answer this question correctly, students should know the characteristics of a normal distribution, especially that the variance, and therefore that the standard deviation is positive. In addition, they should be able to distinguish between a normal and a standard normal distribution with respect to their characteristics. Thirty-five percent of the students over-generalized the characteristics of the standard normal distribution to normal distribution (S.D. =1), while 22% distinguished between the two but included negative numbers as possible values of the S.D. It is likely that this error stems from using knowledge acquired previously concerning the square root. It is also reasonable that these students failed to assimilate the S.D. as a mean of the squared deviations (distances) of the variable values from the mean of the distribution.

**Q6. Comparison Among Benchmark Values Associated with a Normal Distribution**

*Rank the values of $\bar{X}$, Q3, Q1, Z=+1, in descending order, for a normal distribution.*

To answer this question, students' should (a) know what each one of these symbols represents, (b) understand the meaning of the measures represented by these symbols, and (c) be able to use relevant knowledge about the normal distribution and know how these measures relate to it. In addition, they should be able to transform all these values to a common scale (Z) if they cannot arrive at their order without such a transformation. Almost all students were able to correctly rank $\bar{X}$ and Q1 although many (45%) mistakenly ranked Z=+1 and Q3. Students who erred in this question were probably unable to integrate their knowledge about these measures and failed to properly apply it to the normal distribution.

**Q7. Measures of Central Tendency and Dispersion**

*When a constant is added to each value of a series of numbers, which of the following is changed?*

*1. The mean 2. The range 3. The standard deviation 4. All the above*

A correct answer to this question means that students understand the definition of these measures and are able to apply that knowledge to examine how these measures change as a result of certain manipulations. The results indicated that about 31% of the students mistakenly selected option 4 as an answer. This result can be attributed to the tendency of students, particularly those who are not mathematically oriented, to treat statistical measures as raw numbers. Furthermore, students tend to erroneously generalize facts that are correct for certain measures to other measures without critically evaluating their decisions. One remedial practice in this case is to require students to test their answer by computing these measures for the original and the manipulated series and then comparing the resulting values.

**Q8. Scientific Inquiry**

*What do we mean by a "scientific argument"?*

*1. A logical argument*

*2. An argument that is acceptable by experts in the domain of interest*

*3. An argument based on observations*

*4. An argument that allows retrieving the research findings*

Answering this question correctly requires students to know and understand the definition of a scientific argument. The difficulty level of this question is medium (item difficulty =56.8%). Option 4 attracted 39.5% of the respondents. It seems that the concepts "observations" in option

BEST COPY AVAILABLE

3 and findings in option 4 were confused, which may have caused the erroneous answer. It should be noted that the topic "scientific inquiry" was included in the required reading although students attend no lectures about it. Therefore, it is probable that many students did not read the assigned material and answered this question by guessing.

## Summary and Conclusions

The main purpose of this study was to explore the errors committed by students in an introductory statistics test. The vast majority of the questions on the test focused on examining students' conceptual understanding of statistical concepts and procedures and on their ability to apply the knowledge they assumed to have acquired to solve the given tasks. Content analysis of the responses to the test questions yielded several types of errors, including confusing concepts, using irrelevant procedures, using partial information, misinterpreting descriptive information, and making computational errors.

The majority of errors found in this study are not a function of computational difficulties, nor they a function of carelessness and responding without engaging in the task. These results confirm similar findings reported by Mevarech (1983) and more recently by Cohen et al. (1996). The three studies conclude that students' errors may be rooted in difficulties to properly apply their knowledge concerning the different concepts and statistical measures. Furthermore, a potential source of many of the errors detected in this study may result from assimilating statistical concepts into inappropriate schema already existing. For example, errors associated with question 7 indicates that students used mean scores and variance as ordinary numbers and forgot that they are measures of central tendency and dispersion.

Allwood (1984) also argued that even when knowledge resources are available, that availability is insufficient to warrant a successful solution if students are unable to use them.

Moreover, even when the combination of available knowledge resources and the ability to apply them are accessible, the possibility to successfully solve the problem depends on the interaction between them and the nature of the task. Allwood's argument receives some support in this study. For example, question 4 demonstrate how the nature of the problem, among other things, determines the probability of arrivingat successful solutions. Question 4 requires that knowledge about several measures be available to the solver and that s/he be able to use them appropriately. Nonetheless, the complexity of the task imposes the additional difficulty of relating relevant knowledge from different sources in a way that is expedient to the correct solution. In other words, in order to be as successful a possible in solving the problem, the solver should be able to utilize knowledge resources in a way that is advantageous for obtaining the correct solution. Overall, and consistent with Allwood's (1984) view, our results indicate that the outcome of an individual's problem solving activites is determined by a combination of the knowledge resources and the way these resources are utiized in interaction with the problem itself.

Although this study can be viewed as an initial step towards understanding students' reasoning in applying statistical knowledge, it provides teachers of introductory statistics courses with important additional information concerning the content and type of statistical errors their students are likely to commit, and their possible sources. Teachers' knowledge of error patterns in problem solving can promote positive outcomes. An appropriate student-error analysis can lead not only to accurate feedback on the source of those errors, but it can also reveal the students' underlying conceptual misunderstandings and/or lack of proper problem-solving strategies. The error analysis approach to students' thinking offers the best chance currently available for identifying problems and improving teaching. This approach can ultimately become the basis for further instructional planning on the part of the teacher, and curriculum design on the part of the

educational system.

The results of this study suggest that further research can profitably be directed towards understanding students' thinking by diagnosing their errors and misconceptions. It would be more beneficial and meaningful to follow this exploratory study by a series of in-depth investigations, each focusing on a specific statistical topic. Further research is also needed to shed light on the sources of students' errors to inform decisions concerning remedial instruction.

References

Allwood, C. M. (1984). Error detection processes in statistical problem solving. Cognitive Science, 8, 413-437.

Auizmendi, E. (1991, April). Factors related to attitudes toward statistics: A study with a Spanish sample. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Babbitt, C. B. (Oct. 1990). Error patterns in problem solving. Paper presented at the International Conference of the Council for Learning Disabilities. Austin, TX.

Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. Cognitive Science, 2, 155-192.

Cohen, S., Smith, G., Chehile, A. R., Burn, G., & Tsai, F. (1996). Identifying impediments to learning probability and statistics from an assessment instructional software. Journal of Educational and behavioral Statistics, 21, 35-54.

Del Vecchio, A. (1995). A psychological model of introductory statistics course completion. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Feinberg, F., &. Halprin, S. (1978). Affective and cognitive correlates of course performance in introductory statistics. Journal of Experimental Education, 46, 11-18.

Gal, L. B., & Ginsburg, L. (1994). The rule of beliefs and attitudes in learning statistics: Toward an assessment framework. Journal of Statistics Education 2.

Glencross, M. J., & Cherian, V. I. (1992). Attitudes toward applied statistics of postgraduate students in Transco. Psychological Reports, 70, 67-75.

Herman, E. W. (August, 1997). Statistical content errors for students in an educational psychology course. Paper presented at the Annual Meeting of the American Psychological Association. Chicago, IL.

Jitendra, K. A., & Kameenui, J. E. (1993). Experts' and Novices Error patterns in solving part-whole mathematical word problems. Journal of Educational Research, 90, 42-51.

Mevarech, R. Z. (1983). A deep structural model of students' statistical misconceptions. Educational studies in mathematics, 14, 415-429.

Nasser, F. (1989). Effects of sex, test anxiety and item sequence on performance on diagnostic test in exponents. Unpublished master's thesis, Tel Aviv University, Tel Aviv.

Nasser ,F. (July, 1998). Attitudes toward statistics and statistics anxiety among college students: Structure and relationship to prior mathematics experience and performance in introductory statistics course. Paper presented at the annual meeting of the Stress and Anxiety Society (STAR), Turkey, Istanbul.

Nasser ,F. (July, 1999). Statistics anxiety: Structure and relationship to mathematics anxiety, attitudes about mathematics and statistics, motivation, and achievement in an introductory statistics course. Paper presented at the annual meeting of the Stress and Anxiety Society (STAR), Cracow, Poland.

O'Connell, A. A. (November, 1993). Investigating the relationship between conceptual and procedural errors in the domain of probability problem solving. Paper presented at the annual meeting of the Mid-South Educational Research Association. New Orleans, LA.

Shaughnessey, J. M. (1993). Connecting research to teaching probability and statistics. Mathematics Teacher, 86, 244-248.

Thompson, A. P., & Smith, L. M. (1982). Conceptual, computational, and attitudinal correlates of student performance in introductory statistics. Australian Psychologist, 17, 191-197.

Vidal-Madjar, A. ( 1978). Teaching mathematics and statistics to adults who are keen on psychology. Educational studies in Mathematics, 9, 381-390.

Vikali, M. (1985). Children's thinking in arithmetic word problem solving. The Journal of Experimental Education, 53, 106-113.

Watts, D. G., (1991). Why introductory statistics difficult to learn? And what we can do to make it easier?. The American Statistician, 45, 290-405.

Wisenbaker, J. M., & Scott, J. S. (1997). Modeling aspects of students' attitudes and achievement in introductory statistics course. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

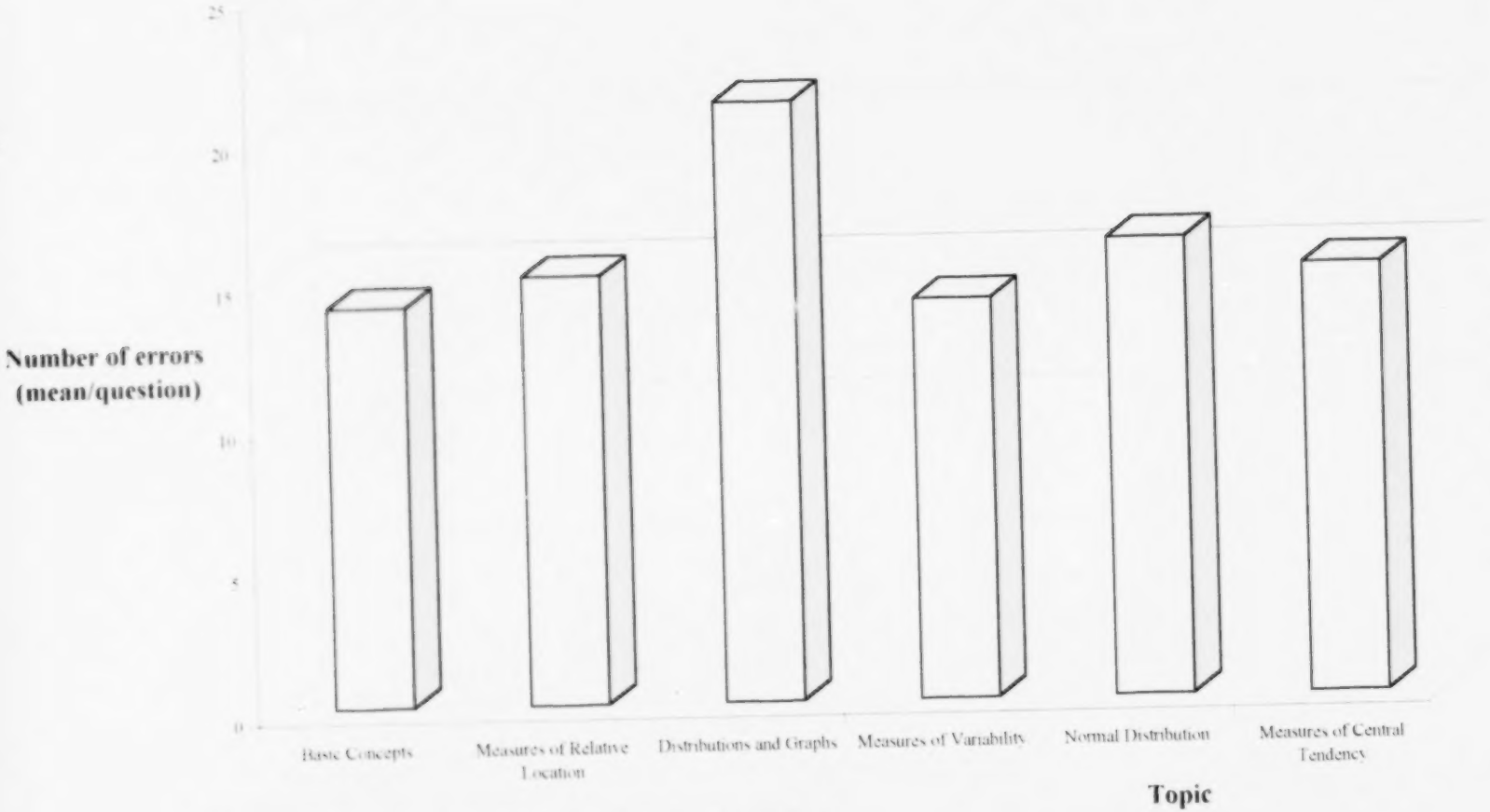


**Figure 1.** **Number of errors per topic (mean/question).**

Table 1

The Achievement Test: Summary of the Statistical Topics, Subtopics, Task Descriptions, Frequent Errors, and Item Difficulty Indices.

| Topic | Subtopic | Task | Frequent Errors | Item No | Item Difficulty (% Correct) |
|---|---|---|---|---|---|
| • Basic concepts in statistics and research methods | • Hypotheses | • Identification of research hypotheses from a research synopsis | • Confusing hypothesis, theory and research problem | 1 | 92.1 |
| | • Variables | • Defining variables (nominal and operational) | • Confusing nominal and operational defintions of a given variable | 2 | 53.1 |
| | | • Distinguishing between a variable and its values (types) | • Confusing a variable and its values | 49 | 82.7 |
| | • Measurement direction | • Identifying dependent and independent variables | • Confusing dependent and independent variables | 3<br>4<br>5 | 93.8<br>85.2<br>67.9 |
| | • Measurement scale (level) | • Proposing an appropriate scale for measuring a given variable<br>**or**<br>• Selecting an appropriate scale for measuring a given variable | • Selecting an inappropriate scale for measuring a given variable<br>**or**<br>• making an erroneous match between a given scale and a selected variable | 21<br>35<br>36<br>37<br>38 | 85.2<br>72.8<br>92.6<br>74.1<br>79.0 |
| | • Scientific Inquiry | • Understanding the meaning of scientific arguments.<br>• Identifying principles of scientific inquiry | • Erroneous comprehension of the scientific argument<br>**or**<br>• Misunderstanding of a scientific principle (the generalization principle) | 43<br>44 | 56.8<br>62.5 |

Table 1 (continued)

| Topic | Subtopic | Task | Frequent Errors | Item (% Correct) | Item Difficulty No. |
|---|---|---|---|---|---|
| • Measures of relative location | • Z-scores | • Comparison of relative locations on different distributions | • Referring to the index of variability instead of Z-score | 6 | 53.1 |
| | | | • Computational errors (the difference between two negative numbers | 22 | 86.4 |
| | | | • Misinterpreting the value of the standard deviation (Lower S.D. interpreted to mean higher rank on the variable of interest) | 26 | 82.7 |
| | • Percentiles | • Interpretation of the percentile | • Calculatiing the mean of two percentiles instead of interpreting the relevant percentile given in the question (students feel obligated to use all the information given in the question even when part of this information is irrelevant. | 8 | 84.0 |
| • Distributions and graphs | • Symmetric distribution | • Determining the shape of the distribution based on a given frequency table | • Confusing negative and positive tails | 7<br>14 | 58.1<br>81.5 |
| | • The relation between the shape of the distribution and measures of central tendency | • Identifying the appropriate measures of central tendency for describing a given distribution | • Selecting inappropriate measures of central tendency to describe a given distribution (the frequent errors reflect confusion of the mean and the median) | 8<br>17 | 71.6<br>37.4 |
| | | • Determining the shape of the distribution on the basis of given measures of central tendency | • A symmetric distribution is described by the median | 10 | 84.1 |

Table 1 (continued)

| Topic | Subtopic | Task | Frequent Errors | Item N0. | Item Difficult (% Correct) |
|---|---|---|---|---|---|
| • Measures of variability (Dispersion) | • Variance and standard deviation | • Determining the value of the standard deviation of a given distribution (a series of identical values) | • Generalizing the characteristics of the standard normal distribution to any symmetric distribution (S.D.=1) | 11 | 87.7 |
| | | • Interpreting the standard deviation | • S.D.= 0 means that the number of negative and positive values equal zero | 31 | 79.0 |
| | | • Relating the standard deviation to the Z-score | • Computational errors (substracting two negative numbers) | 13 | 82.7 |
| | • Range and Inter-quartile range | • Relating Q to the 50% of the cases at the center of the distribution | • Lack of attempt error (there is insufficient information to answer the question) | 12<br>16 | 60.5<br>59.3 |
| | | • Obtaining the inter-quartile range from given data | • Obtaining Q by calculating the the difference between Q3 and the Median | 15 | 82.7 |
| | | • Selecting the appropriate dispersion measure(s) for describing a given distribution | • Selection of the S.D. for describing the dispersion of an asymmetric distribution | 29 | 74.1 |
| | • Index of variability | • Comparison of two groups regarding their homogeneity with regard to a given variable | • Small S.D. interpreted as a lower level of the variable instead of a measure of homogeneity and use of partial information S.D. instead of C.V. | 25 | 74.1 |

Table 1 (continued)

| Topic | Subtopic | Task | Frequent errors | Item No. | Item Difficulty (% Correct) |
|---|---|---|---|---|---|
| • Normal distribution | • Meaning of Z and its relation to standard deviation | • Obtaining the standard deviation given the range | • Ignoring the given data and giving the answer of an in-class example | 18 | 70.4 |
| | • Characteristics of the normal distribution | • Distinguish between normal and standard normal distribution | • Over-generalization of the characteristics of the standard normal distribution to other symmetric and asymmetric distributions | 20 | 37.0 |
| | | • Determining the probability associated with a given value of the variable of interest | • Instead of looking at 0 +1 S.D. they looked at 0± S.D. | 23 | 88.9 |
| | | • The relation between X and Z | • Misunderstanding the meaning of Z (when Z=0 then X=0) | 24 | 91.4 |
| | | • Determining the locations of benchmark values on the Z | • In a normal distribution $Q_3$>Z=+1 | 30 | 54.3 |

Table 1 (continued)

| Topic | Subtopic | Task | Frequent errors | Item No. | Item Difficulty (% Correct) |
|---|---|---|---|---|---|
| • Measures of central tendency | • Mode | • Selecting the appropriate measure(s) of central tendency based on a summary table | • Confusion between mode and frequency (in Hebrew these two words are very close in sound and spelling) | 27 | 74.1 |
| | • Median | • As above | • Selection of inappropriate measure (s) of central tendency for describing a given distribution (use of the mean for describing an asymmetric distribution) | 28 | 63.0 |
| | • Mean | • Understanding characteristics of the measures of central tendency | • All measures of central tendency change as a result of adding a constant to all values of a statistical series | 33 | 60.5 |
| | • The effect of manipulation on measures of central tendency | • Obtaining and comparing the the mean and median of two distributions | • Only the means of the two series are equal (answer is that both the mean and the median are equal) | 34 | 88.9 |